

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:46:36 ; Search time 2644 Seconds

(without alignments)
6098.188 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 taccctggggccagatcc.....ttccgctactagtagccgcg 372

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ses:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ses:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sv:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214.4	57.6	1181	10	RATLIGFZ
2	214.4	57.6	185148	2	AC136382
3	188.8	50.8	11363	10	MMILGF
4	188.8	50.8	192843	10	AL607124
5	98.2	26.4	1569	10	BC013345
6	76.4	20.5	1500	6	AX401932
7	76.4	20.5	1500	6	AX827271
8	76.4	20.5	1500	10	RATIGFB
9	76.4	20.5	5001	6	AX163782
10	76.4	20.5	5001	10	RATIGFBA
11	64	17.2	3886	6	AY095345
12	63.8	17.2	6128	6	AX409747
13	63.8	17.2	6128	9	HUMIGFBP1A
14	63.8	17.2	6128	11	G19994
15	63.8	17.2	6480	9	HUMIGFBP1
16	63.8	17.2	9082	9	AY434089
17	63.8	17.2	69887	9	AC091524
18	62.2	16.7	141539	9	AC146152
19	62.2	16.7	182079	2	AC146117
20	60.8	16.3	448	11	G67139
21	55.2	14.8	194	6	A57715
22	55.2	14.8	194	6	AR175909
23	55.2	14.8	13011	6	E14395
24	55.2	14.8	13011	6	AX827302
25	55.2	14.8	13011	10	RMLPG
26	54.4	14.6	895	9	HUMIGFBP1
27	52	14.0	539	11	G67173
28	47.4	12.7	2717	10	RATPKRL1
29	47.4	12.7	231241	2	AC097039
30	47.2	12.7	199854	2	AC132327
31	44.6	12.0	18489	9	AL442125
32	44.6	12.0	256781	2	AC097952
33	44.6	12.0	259329	2	AC097950
34	44.4	11.9	1480	10	RATIGFBP1
35	44.2	11.9	206515	2	AC140332
36	42.6	11.5	43295	2	AC006177
37	42.6	11.5	53370	9	AL592071
38	42.6	11.5	186314	10	AC003694
39	42.4	11.4	185452	2	AC114566
40	42.4	11.4	187942	2	AC116502
41	41.4	11.1	200340	9	AC100797
42	41	11.0	81704	9	AC110299
43	41	11.0	151700	9	AC133528
44	41	11.0	182444	2	AC116481
45	40.6	10.9	125020	9	AF429315

ALIGNMENTS

RESULT 1
LOCUS RATLIGFZ
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION M84484.1
VERSION M84484.1 GI:204927
KEYWORDS insulin-like growth factor.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1181)
REFERENCE
Unterman,T.G., Iacson,R.G., McGary,E., Whalen,C. and Goswami,R.G.
JOURNAL Biochem. Biophys. Res. Commun. (1991) In press

COMMENT Original source text: Rattus norvegicus (strain Sprague-Dawley)
male adult liver DNA.

FEATURES
source
1. 1181
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

Query Match 57.6%; Score 214.4; DB 10; Length 1181;
Best Local Similarity 97.3%; Pred. No. 1.2e-51;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 144 TGGCCCATGTGCACAGCAAAACAACTATTATTGACACGGGATCCTAGACGCTGCC 203
DB 813 TGTGGCGAGCTCACAGCAAAACAACTATTATTGACACGGGATCCTAGACGCTGCC 872
QY 204 CTGACATCATTTACCGGCTGCGAGCGAGCCCTTATTAAGCCCTGGATGGCCAG 263
DB 873 CTGACATCATTTACCGGCTGCGAGCGAGCCCTTATTAAGCCCTGGATGGCCAG 932
QY 264 CCAGCATGTCCTACCTCCCGCCGAGACACAAACCCAGGAGCATTTGAACTGCACACG 323
DB 933 CCAGCATGTCCTACCTCCCGCCGAGACACAAACCCAGGAGCATTTGAACTGCACACG 992
QY 334 CCATTTGCCGACAGAGCTGTGTACCCACCTTCCTGCTACTAGCTA 367
DB 993 CCATTTGCCGACAGAGCTGTGTACCCACCTTCCTGCTACTAGCTA 1036

RESULT 2
AC136382 185148 bp DNA linear HTG 01-NOV-2002
LOCUS Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
DEFINITION *** 63 unordered pieces.
AC136382
AC136382.1 GI:24462257
HTG; HTGS_PHASE1.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus;
1 (bases 1 to 185148)
Muzny,D.Marie, Metzker,M.Lee, Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryan,N., Bunay,C., Burch,P., Burrell,K., Calderon,B.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,D.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Daviela,M.L., Davis,C., Davy-Carrroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazier,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunnarane,P., Healand,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Huik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Kapatich,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kovis,C., Kraft,C.L., Lebow,H., Lavan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Louisedge,H., Lozago,R.J., Lu,X., Ma,J.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mareshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangun,A.,
Mangun,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwokilemeh,O., Okwou,G.,
Olatunmbesogun,A., Pal,S., Parks,K., Patelnak,S., Paul,H.,
Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Polindexter,A.,
Popovic,D., Primus,E., Pu,L.-L., Puzo,M., Quiroz,J., Rachlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shatman,S., Shen,H., Shetty,J., Shvartbeyn,A.,
Sison,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
Svaltek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D.,
Waidron,L., Walker,B., Wang,D., Wang,Q., Wang,S., Warren,D.,
Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,X.,
Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 185148)
Rat Genome Sequencing Consortium.
Submitted (01-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: KDG5
Center clone name: CH230-97018

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 135613 bases at least Q40
Consensus quality: 140849 bases at least Q30
Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1010: contig of 1010 bp in length
1011 1110: gap of unknown length
1111 2608: contig of 1498 bp in length
2609 2708: gap of unknown length
2709 3915: contig of 1207 bp in length
3916 4015: gap of unknown length
4016 5547: contig of 1532 bp in length
5548 5647: gap of unknown length
5648 7119: contig of 1472 bp in length
7120 7220: gap of unknown length
7221 8885: contig of 1666 bp in length
8886 10258: gap of unknown length
10259 10358: contig of 1273 bp in length
10359 11627: gap of unknown length
11628 11726: contig of 1268 bp in length
11727 11726: gap of unknown length

```

11727 13488: contig of 1762 bp in length
13489 13588: gap of unknown length
13589 15132: contig of 1534 bp in length
15123 15222: gap of unknown length
15223 16946: contig of 1724 bp in length
16946 17046: gap of unknown length
17046 18164: contig of 1118 bp in length
18164 18264: gap of unknown length
18264 19678: contig of 1414 bp in length
19678 19778: gap of unknown length
19778 21687: contig of 1909 bp in length
21687 21787: gap of unknown length
21788 22892: contig of 1105 bp in length
22892 24336: gap of unknown length
24336 24436: contig of 1344 bp in length
24436 26617: gap of unknown length
26617 26717: contig of 2181 bp in length
26717 28577: contig of 1860 bp in length
28577 28677: gap of unknown length
28677 30823: contig of 2146 bp in length
30823 30923: gap of unknown length
30923 32266: contig of 1343 bp in length
32266 32367: gap of unknown length
32367 33894: contig of 1528 bp in length
33894 33994: gap of unknown length
33994 35373: contig of 1379 bp in length
35373 35473: gap of unknown length
35473 37295: contig of 1822 bp in length
37295 37395: gap of unknown length
37395 38420: contig of 1025 bp in length
38420 38521: gap of unknown length
38521 39974: contig of 1454 bp in length
39974 40074: gap of unknown length
40074 42283: contig of 2209 bp in length
42283 42384: gap of unknown length
42384 43721: contig of 1338 bp in length
43721 43821: gap of unknown length
43821 46538: contig of 2717 bp in length
46538 46638: gap of unknown length
46638 48621: contig of 1983 bp in length
48621 48722: gap of unknown length
48722 50602: contig of 1881 bp in length
50602 50702: gap of unknown length
50702 53841: contig of 3139 bp in length
53841 53942: gap of unknown length
53942 56195: contig of 2254 bp in length
56195 56295: gap of unknown length
56295 58949: contig of 2654 bp in length
58949 59050: gap of unknown length
59050 61032: contig of 1983 bp in length
61032 61132: gap of unknown length
61132 62942: contig of 1810 bp in length
62942 63043: gap of unknown length
63043 66045: contig of 3003 bp in length
66045 66145: gap of unknown length
66145 68197: contig of 2052 bp in length
68197 68298: gap of unknown length
68298 72235: contig of 3938 bp in length
72235 72336: gap of unknown length
72336 75803: contig of 3468 bp in length
75803 75904: gap of unknown length
75904 78355: contig of 2452 bp in length
78355 78456: gap of unknown length
78456 81458: contig of 3003 bp in length
81458 84673: contig of 3115 bp in length
84673 84774: gap of unknown length
84774 87739: contig of 2966 bp in length
87739 87839: gap of unknown length
87839 91535: contig of 3696 bp in length
91535 91636: gap of unknown length
91636 95522: contig of 3887 bp in length

```

```

* 95523 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98830: gap of unknown length
* 98831 101618: contig of 2788 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117630: contig of 3966 bp in length
* 117631 117731: gap of unknown length
* 117731 122503: contig of 4772 bp in length
* 122503 122602: contig of 4772 bp in length
* 122602 122603: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126644 126744: gap of unknown length

```

Query Match 57.6%; Score 214.4; DB 2; Length 185148;
 Best Local Similarity 97.3%; Pred. No. 1.6e-51;
 Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 144 TGGCCCATGTCACAGCAAAACAACTATTATTGACACGGGATCTTAGCAGCGTGC 203
Db 99243 TGGCGGAGCTCAGACAAACAACTATTATTGACACGGGATCTTAGCAGCGTGC 99302
QY 204 CTGACATATTAAACCCGTCGTCGCGAGCCGCTTCTAAGGCCCTGGGTATGCGCAG 263
Db 99303 CTGACATATTAAACCCGTCGTCGCGAGCCGCTTCTAAGGCCCTGGGTATGCGCAG 99362
QY 264 CCAGCATGTCACCTCCCGCCGAGACACAAACCCAGCAGATTGACATGTCACACG 323
Db 99363 CCAGCATGTCACCTCCCGCCGAGACACAAACCCAGCAGATTGACATGTCACACG 99422
QY 324 CCATCTGCCAGAGAGCTGTGACCACTCTCGCTACTAGCTA 367
Db 99423 CCATCTGCCAGAGAGCTGTGACCACTCTCGCTACTAGCTA 99466

```

```

RESULT 3
LOCUS MMLIGF 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493.1 GI:52699
VERSION X67493.1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1363)
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohr,K.L., Bucan,M. and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
JOURNAL Hepatology 19 (3), 656-665 (1994)
MEDLINE 94164648
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohr,K.L., Waddell,J.R. and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
JOURNAL Nucleic Acids Res.
AUTHORS Taub,R.A.
TITLE Direct Submissio
JOURNAL Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source 1..1363
/organism="Mus musculus"

```

```

/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
misc_feature
692..706
/note="insulin-responsive element"
CAAT_signal
720..741
/note="CAAT box with APF, HNF and NF-E1 consensus
sequences"
764..768
792..1336
/number=1
TATA_signal
792
/note="mRNA cap site"
misc_feature
964..1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="SWISS-PROT:P47876"
/translation="MPEFLTVSMNPFILSFQIGVAGAPQPMHCACTAERIGLCP
PVPAACPRISPAAGCGCCPTCAIPMGAAAGVATACAGISCRALPGEPRPLHLATRG
QASLPFAAPATSTLFSQHE"
ORIGIN

```

```

Query Match      50.8%; Score 188.8; DB 10; Length 1363;
Best Local Similarity 90.2%; Pred. No. 3.6e-44;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 144 TGGCCCATGTGCACAGCAAACTTATTGTAACACGGGATCTTACACGCTGCC 203
DB 667 TGTGTAGAGCTCACAGCAAACTTATTGTAACACGGGATCTTACACGCTGCC 726
QY 204 CTGCAATCATTAAACCGTGTGCGGACGAGCCCTTATAGAGCCCTGGGTATGGCCAG 263
DB 727 CTGCAATCATTAAACCGTGTGCGGACGAGCCCTTATAGAGCTTGGGTATGGCCAG 786
QY 264 CCAGCATGTGTCACCTGCGCGGACGAGCAAAACCCAGGAGCATTTGAACATGACACG 323
DB 787 CCAGCATGTGTCACCTGCGCGGAGACACACACCCAGGAGCATTTGAACATGACACG 846
QY 324 CCATCTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 847 CCGTCTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 890

```

```

RESULT 4
AL607124      192843 bp  DNA      linear  ROD 11-APR-2002
LOCUS        Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
DEFINITION
ACCESSION    AL607124
VERSION      AL607124.15  GI:20145926
KEYWORDS     HTG.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1
AUTHORS      Oliver, K.
TITLE        Direct Submision
JOURNAL      Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
              Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              clonerequest@sanger.ac.uk
              humquerry@sanger.ac.uk
              On Apr 12, 2002 this sequence version replaced gi:19847866.
COMMENT      During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is from the RP23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: PBAC3.6.

```

FEATURES
source
Location/Qualifiers
1..192843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-20C9"
/clone_lib="RP23-20C9"
ORIGIN

```

```

Query Match      50.8%; Score 188.8; DB 10; Length 192843;
Best Local Similarity 90.2%; Pred. No. 4.6e-44;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 144 TGGCCCATGTGCACAGCAAACTTATTGTAACACGGGATCTTACACGCTGCC 203
DB 2419 TGTGTAGAGCTCACAGCAAACTTATTGTAACACGGGATCTTACACGCTGCC 2478
QY 204 CTGCAATCATTAAACCGTGTGCGGACGAGCCCTTATAGAGCCCTGGGTATGGCCAG 263
DB 2479 CTGCAATCATTAAACCGTGTGCGGACGAGCCCTTATAGAGCTTGGGTATGGCCAG 2538
QY 264 CCAGCATGTGTCACCTGCGCGGACGAGCAAAACCCAGGAGCATTTGAACATGACACG 323
DB 2539 CCAGCATGTGTCACCTGCGCGGAGACACACACCCAGGAGCATTTGAACATGACACG 2538
QY 324 CCATCTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 2599 CCGTCTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 2642

```

```

RESULT 5
BC013345      1569 bp  mRNA      linear  ROD 12-NOV-2003
LOCUS        Mus musculus insulin-like growth factor binding protein 1, mRNA
DEFINITION   (CDNA clone MGC:14075 IMAGE:4161889), complete cds.
ACCESSION    BC013345
VERSION      BC013345.1  GI:15426482
KEYWORDS     MGC.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, J., Shenmen, C.M., Schuler, G.D.,
              Altschuler, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
              Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Schaeetz, T.E., Brownstein, M.J., Utsch, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

```

ABRAMSON, R.D., MULLAHY, S.J., BOSEK, S.A., MCKEY, P.J.,
MCKENNA, K.J., MALEK, J.A., GUNARATNE, P.H., RICHARDS, S.,
MORLEY, K.C., HALE, S., GARCIA, A.M., GAY, L.J., HULYK, S.W.,
VILLALON, D.K., MUZYNY, D.M., SODERGREN, E.J., LU, X., GIBBS, R.A.,
FAHEY, J., HELTON, E., KETTEMAN, M., MADAN, A., RODRIGUES, S.,
SANCHEZ, A., WHITING, M., MADAN, A., YOUNG, A.C., SHEVCHENKO, Y.,
BOUFFARD, G.G., RODRIGUEZ, R.W., TOUCHMAN, J.W., GREEN, E.D.,
DICKSON, M.C., RODRIGUEZ, A.C., GRIMWOOD, J., SCHMUTZ, J., MYERS, R.M.,
BUTTERFIELD, Y.S., KRZYWINSKI, M.I., SKALSKA, U., SMALUS, D.E.,
SCHNECH, A., SCHEIN, J.E., JONES, S.J. and MARRA, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
22388257
12477932
2 (bases 1 to 1569)
Strausberg, R.
Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

REMARK
COMMENT
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://imgc.lnl.gov>
Series: IRAX Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1. 1569
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:14075 IMAGE:416189"
/issue_type="Liver, normal, 5 month old male mouse."
/clone_id="NCI CGAP_L19"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1. 1569
/gene="Igf1bp1"
/note="synonym: IGFBP-1"
/db_xref="LOCUSID:16006"
/db_xref="MGI:96436"
204..1022
/codon_start=1
/product="Insulin-like growth factor binding protein 1"
/protein_id="AA13345.1"
/db_xref="GI:15426483"
/db_xref="LOCUSID:16006"
/translators="MPFLVVSMPFLILISFOIGVAGAPQWCACTAERGLCP
PVPASPEISRPAGCGCCPFCAGACGVAARACAGISCSALGEPPLALTRG
OQACVPEPAAPATSLSSQHEBAKAAVVASDELSSPTEBOLDSFHLMAFSD
OPTLMAISTYSRARAIADLKWKVPCORELYKLERLAAQOQAGDEIFYFYLPN
CKNGFYHSGKQCTSLDGEAGLCMCVYPPMSKKIPESLERGPBNHQYFNNVN"
231..527
/note="1B; Region: Insulin growth factor-binding protein
homologues"
/db_xref="CDD:smart00121"
768..995
/note="thyroglobulin_1; Region: Thyroglobulin type-1

misc_feature
misc_feature
misc_feature

repeat. Thyroglobulin type 1 repeats are thought to be
involved in the control of proteolytic degradation. The
domain usually contains six conserved cysteines. These
form three disulphide bridges. Cysteines 1 pairs with 2, 3
with 4 and 5 with 6"
/db_xref="CDD:pfam00086"
/db_xref="CDD:pfam00086"

ORIGIN
Query Match 26.4%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 1e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 257 TGCGCAGCCAGCATGTGCTCCAGTCCCGCCGAGACACAAACCCAGCATTAAACACTG 316
DB 20 TGGGACGACGATGTGCTCCAGTCCCGCCGAGACACACACCCAGCATTAAACACTG 79

QY 317 CACACGGCCGATGTGCTCCAGTCCCGCCGAGACACACACCCAGCATTAAACACTG 367
DB 80 CACACGGCCGATGTGCTCCAGTCCCGCCGAGACACACACCCAGCATTAAACACTG 130

RESULT 6
AX401932 1500 bp DNA linear PAT 06-JUN-2002
LOCUS
DEFINITION
Sequence 1608 from Patent WO210453.
ACCESSION
AX401932
VERSION
AX401932.1 GI:21338112
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1
Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and
Elaasoff, M.R.
Molecular toxicology modeling
Patent: WO 0210453-A 1608 07-FEB-2002;
Gene Logic, Inc. (US)
Location/Qualifiers
1. 1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/Genbank Accession No. NM_013144"

ORIGIN
Query Match 20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.3e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCAGCAGGATTTGAACACTGCACACGCCCATGTGCCAGAGAGCTGTGACAC 349
DB 8 CACAAACCCAGCAGGATTTGAACACTGCACACGCCCATGTGCCAGAGAGCTGTGACAC 67

QY 350 CACTTCCGCTACTAGCTA 367
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 7
AX827271 1500 bp DNA linear PAT 12-DEC-2003
LOCUS
DEFINITION
Sequence 5 from Patent EP1344834.
ACCESSION
AX827271
VERSION
AX827271.1 GI:39837360
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1

AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
F. HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers
source 1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN
Query Match 20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.3e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 230 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 349
|||||
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
|||||

QY 350 CACTTCGGCTACTACTA 367
|||||
DB 68 CACTTCGGCTACTACTA 85
|||||

RESULT 8
RATIGFB
LOCUS Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.
DEFINITION M58634.1 GI:204732
ACCESSION M58634.1
VERSION M58634.1
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulin-like growth factor-binding protein 1
JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE 91141487
PUBMED 1705004
COMMENT Original source text: Rat, cDNA to mRNA.
FEATURES Location/Qualifiers
source 1..1500
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AA41380.1"
/db_xref="GI:204732"
/translation="MPEFLTVSWPFLILSFQYRVVAGAPQPMHCAPCTARLELCP
VPASCEPISRPAGCCPTCALPLGAGAGVATACAGLSCRLPGRPRHLATRG
GACVLEAPAPATSLSGSHBEAKAASDELAESEPMREBOLDSFHLMAISRED
QPLIMASTYSSMRAREITDILKKKEKCEQELVYERLMAAQKAGDELYKTYLPN
CNKGFFHSKCESTSLDEAGLCWCVYWSKSKIPLSLRGTGDPCHQYFNQV"

ORIGIN
sig_peptide 160..234
mat_peptide 235..975
/product="IGF binding protein-1"

QY 230 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 349
|||||
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 67
|||||
QY 350 CACTTCGGCTACTACTA 367
|||||
DB 68 CACTTCGGCTACTACTA 85
|||||

RESULT 9
AX163782
LOCUS Sequence 46 from Patent WO0138579.
DEFINITION AX163782
ACCESSION AX163782
VERSION AX163782.1 GI:14544878
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Gould-Rothberg, B.E., Dipippo, V.A., Rameeh, T.M. and Gerweil, R.W.
TITLE Method of identifying toxic agents using naaid-induced differential
JOURNAL gene expression in liver
Patent: WO 0138579-A 46 31-MAY-2001;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..5001
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN
Query Match 20.5%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 2.4e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 230 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 349
|||||
DB 1 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCAC 60
|||||

QY 350 CACTTCGGCTACTACTA 367
|||||
DB 61 CACTTCGGCTACTACTA 78
|||||

RESULT 10
RATIGFBA
LOCUS Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,
DEFINITION complete cds.
ACCESSION L22979
VERSION L22979.1 GI:1098472
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 5001)
AUTHORS Lacsos, R., Oehler, D., Yang, E., Goswami, R. and Unterman, T.
TITLE Dideoxy sequencing and structural analysis of the rat insulin-like
JOURNAL growth factor binding protein-1 gene
Biochim. Biophys. Acta 1218 (1), 95-98 (1994)
MEDLINE 94250701
PUBMED 7514892
COMMENT On Nov 30, 1995 this sequence version replaced gi:385167.
Original source text: Rattus norvegicus (strain Sprague-Dawley)
DNA.
FEATURES Location/Qualifiers
source 1..5001
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"

/db_xref="taxon:10116"
 /sex="male"
 /issue_type="liver"
 /dev_stage="adult"
 /issue_1ib="Clontech EMBL-3 SP6/T7"
 60..525
 /number=1
 60..152
 153..4221
 /gene="IGFBP-1"
 join(153..525,1850..2034,2874..3002,4090..4221)
 /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein"
 /db_xref="GI:1098473"
 /db_xref="GI:1098473"
 /translation="MPELVVSWPPLILSROVVAGAPQWMCAPCTARLELCP
 PVPASCEISRAGCCGCTCALPLGACGVATPAACAGISCPALGEPRLPHALRG
 OGACVLEPAAPATSLSGSHEBAKAASVEDLAOSPEMTEQLDSFHLMAFRED
 OPLMNAISTYSMRAREITDLKKMKPCORELYKLERLAACQKAGDEIKYFLPN
 CKKNGFYHKKOCETSLDGRAGLCWCYTPMSGKKIPGSLERDGNCHQYFNQV"
 526..1849
 /gene="IGFBP-1"
 /number=1
 1850..2034
 /gene="IGFBP-1"
 /number=2
 2035..2873
 /gene="IGFBP-1"
 /number=2
 2874..3002
 /gene="IGFBP-1"
 /number=3
 3003..4089
 /gene="IGFBP-1"
 /number=3
 4090..4743
 /number=4
 4222..4743
 3'UTR
 ORIGIN
 Query Match 20.5%; Score 76.4; DB 10; Length 5001;
 Best Local Similarity 98.7%; Pred. No. 2.4e-11;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 290 CACAAACCCAGGAGGATTAACACTGACACGCGCCATCTGCCAGAGAGCTGTGACAC 349
 DB 1 CACAAACCCAGGAGGATTAACACTGACACGCGCCATCTGCCAGAGAGCTGTGACAC 60
 QY 350 CACTTCGGCTACTAGCTA 367
 DB 61 CACTTCGGCTACTAGCTA 78
 RESULT 11
 LOCUS AY095345 3886 bp DNA linear PRI 30-DEC-2002
 DEFINITION Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
 ACCESSION AY095345
 VERSION AY095345.1 GI:20853764
 KEYWORDS
 SOURCE
 ORGANISM
 Papio anubis (olive baboon)
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecidae; Papio.
 1 (bases 1 to 3886)
 Kim,J.U., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
 Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
 Regulation of insulin-like growth factor binding protein-1 promoter
 activity by pRHR and HOXA10 in primate endometrial cells
 Biol. Reprod. 68 (1), 24-30 (2003)

PUBMED 12493691
 REFERENCE 2 (bases 1 to 3886)
 AUTHORS Kim,J.U., Jaffe,R.C. and Fazleabas,A.T.
 TITLE Direct Submission
 JOURNAL Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
 University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
 USA
 FEATURES
 source Location/Qualifiers
 1..3886
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /db_xref="taxon:9555"
 <3655..>3886
 /gene="IGFBP-1"
 <3655..>3886
 /gene="IGFBP-1"
 /product="insulin-like growth factor binding protein-1"
 3655..>3886
 /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein-1"
 /protein_id="AAM23273.1"
 /db_xref="GI:20853765"
 /translation="MSFVVARVWLVLILTVQGVTAAPWQCAPCSAEKTLCPV
 PASCEVTRSAAGCGCCPMCALPLGAACGATAR"
 ORIGIN
 Query Match 17.2%; Score 64; DB 9; Length 3886;
 Best Local Similarity 60.6%; Pred. No. 1e-07;
 Matches 149; Conservative 0; Mismatches 80; Indels 17; Gaps 2;
 QY 120 TCAGGAACACGAGAGTCCCGCGCCCATGTACAGCAAGCAACAACTTATTGTA 179
 DB 3331 TCCTCCACCTCGGCTTTCGTAAGGCGCTTGCGCGACATGAAACAACTTATTGTA 3390
 QY 180 ACACGGGAGTCTTACAGACGCTGCCCTGACATCATTTAACCC-----GTGTCGCG 229
 DB 3391 ACACGTAGCTCTTACAGTGTCCCGCGCTGCCAATCATTTAACCTCTGTGCAAGTGGCGCG 3450
 QY 230 AGCCAGCCCTTATTAAGGCGCTGGTATAGGCCAGCCAGCATGATCCTCCGCCCGGAGA 289
 DB 3451 CCTGTGCCCTTATTAAGGCGCGCTGTGTCCAGCAAGATGCGCACCGCATCCATC 3510
 QY 290 CACAAACCCAGGAGGATTAACACTGACACGCGCCATCTGCCAGAGAGCTGTGACAC 349
 DB 3511 CAGCAAGC-----ATGCGCGCGCGCGCGCGCACCTCCAGAGAGACTGGCCAC 3563
 QY 350 CACTTC 355
 DB 3564 CGCTCC 3569
 RESULT 12
 LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
 DEFINITION Sequence 2394 from Patent WO0229103.
 ACCESSION AX409747
 VERSION AX409747.1 GI:21442452
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1
 Alvaras,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 2394 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source Location/Qualifiers
 1..6128
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN /note="EMBL/GenBank Accession No. M74587"

Query Match 17.2%; Score 63.8; DB 6; Length 6128;

Best Local Similarity 63.2%; Pred. No. 1.2e-07; Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Qy 120 TCAGAGAACGACGAGAGTCCCGCCGCGCCATGTCACAGCAAAACAACTATTATTTGA 179
Db 438 TCCTCCCAACGAGCGGTTTGGCGTAGGGCCCTTGCGTCACTAGCAAAACAACTATTATTTGA 497
Qy 180 ACACGGGATCTCTGACAGCTGCGCTGCAATCAATTAACCC-----GTGCTGCGG 229
Db 498 ACACCTACGCTCTGAGCGGCGCGCGCCCAATCAATTAACCTGCTGCAAGTGCGCGG 557
Qy 230 AGCCAGCCCTTCATTAAGCCCGCTGATGCGACGACAGATGTCCTACTGCGCCCGGAGA 289
Db 558 CCTGTGCCCTTTATTAAGGTGGCGGCTGTGTCCAGACGATCGCCACCGCCATCC--- 613
Qy 290 CACAAACCCAGCAGCATTTGAACACTGC-ACACGGCCATCTGCCCCAGAGCTGTGACCA 348
Db 614 ----CATCAGAGGACATCTGCGCGCGCGCGCCGACCCCTCCAGAGAGACTGGCCA 669
Qy 349 CCACTTC 355
Db 670 CCGCTCC 676

RESULT 13

HUMIFBP1A

LOCUS HUMIFBP1A 6128 bp DNA linear PRI 08-NOV-1994
DEFINITION Human insulin-like growth factor binding protein (hIGFBP1) gene,
complete cds.

ACCESSION

M74587

VERSION 1
insulin-like growth factor binding protein.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 6128)
Luthman, R.

Authors Ehrenborg, E., Larsson, C., Stern, I., Jansson, M., Powell, D. R. and
Contiguous localization of the genes encoding human insulin-like
growth factor binding proteins 1 (IGBP1) and 3 (IGBP3) on
chromosome 7

JOURNAL Genomics 12 (3), 497-502 (1992)
MEDLINE 92217971
PubMed 1373120

COMMENT Original source text: Homo sapiens (tissue library: lambda
Charon4A) liver DNA.

FEATURES
Location/Qualifiers

source

1..6128

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="7p13-p12"

/tissue_type="liver"

/tissue_lib="lambda Charon4A"

CAAT_signal

525..529

/gene="IGFBP1"

/note="G00-120-075"

TATA_signal

569..573

/gene="IGFBP1"

join(597..1110,2657..2826,4041..4169,5069..5769)

gene

1110..1110

/gene="IGFBP1"

join(597..1110,2657..2826,4041..4169,5069..5769)

mRNA

597..1110

/product="insulin-like growth factor binding protein 1"

/note="G00-120-075"

exon

597..1110

/gene="IGFBP1"

CDS

/note="G00-120-075"

join(762..1110,2657..2826,4041..4169,5069..5200)

/gene="IGFBP1"

/codon_start=1

/product="insulin-like growth factor binding protein 1"

/protein_id="AA52784.1"

/db_xref="GI:184812"

/db_xref="GDB:G00-120-075"

/translation="MSRPVAVLVLLLTVOGVTAGAPWOCAPCSAEKALCPV
SASCEVTRSGCCGCPMCLPLGAGVATARCAGISGRLRGGQPLHTRGGG
ACVESDASAPHAABAGSPSPSESTETTBEBLDNFIHMAPSEEDHSLMDISTYDS
SKALHVTNIKMKPCRIELRVVESLAKQETISGEISKPYLPNCNKNGYHSRQCS
TSDMEAGLCMCVYPMNKRIIPSPBIIRDPCMCMPNVQN"

762..836

/gene="IGFBP1"

/note="G00-120-075"

join(837..1110,2657..2826,4041..4169,5069..5197)

/gene="IGFBP1"

/product="insulin-like growth factor binding protein 1"

/note="G00-120-075"

1111..2656

/gene="IGFBP1"

/note="G00-120-075"

1405

/gene="IGFBP1"

/note="G00-120-075"

2657..2826

/gene="IGFBP1"

/note="G00-120-075"

2827..4040

/gene="IGFBP1"

/note="G00-120-075"

4041..4169

/gene="IGFBP1"

/note="G00-120-075"

4170..5068

/gene="IGFBP1"

/note="G00-120-075"

4736

/gene="IGFBP1"

/note="G00-120-075"

4780

/gene="IGFBP1"

/note="G00-120-075"

4820

/gene="IGFBP1"

/note="G00-120-075"

5069..5769

/gene="IGFBP1"

/note="G00-120-075"

ORIGIN

Query Match 17.2%; Score 63.8; DB 9; Length 6128;

Best Local Similarity 63.2%; Pred. No. 1.2e-07; Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Qy 120 TCAGAGAACGACGAGAGTCCCGCCGCGCCATGTCACAGCAAAACAACTATTATTTGA 179
Db 438 TCCTCCCAACGAGCGGTTTGGCGTAGGGCCCTTGCGTCACTAGCAAAACAACTATTATTTGA 497
Qy 180 ACACGGGATCTCTGACAGCTGCGCTGCAATCAATTAACCC-----GTGCTGCGG 229
Db 498 ACACCTACGCTCTGAGCGGCGCGCGCCCAATCAATTAACCTGCTGCAAGTGCGCGG 557
Qy 230 AGCCAGCCCTTCATTAAGCCCGCTGATGCGACGACAGATGTCCTACTGCGCCCGAGA 289
Db 558 CCTGTGCCCTTTATTAAGGTGGCGGCTGTGTCCAGACGATCGCCACCGCCATCC--- 613
Qy 290 CACAAACCCAGCAGCATTTGAACACTGC-ACACGGCCATCTGCCCCAGAGCTGTGACCA 348
Db 614 ----CATCAGAGGACATCTGCGCGCGCGCGCCGACCCCTCCAGAGAGACTGGCCA 669
Qy 349 CCACTTC 355

Db 670 CCGCTCC 676

RESULT 14
LOCUS 6128 bp DNA linear STS 28-SEP-1998
DEFINITION 8555752 Eric D. Green Homo sapiens STS genomic, sequence tagged

ACCESSION G19994
VERSION G19994.1 GI:1254693
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Bouffard, G.G., Iyer, L.M., Idol, J.R., Braden, V.V., Cunningham, A.F., Weitzsaeck, L.A., Mohr-Tidwell, R.M., Peluso, D.C., Fulton, R.S., Leckie, M.P. and Green, E.D.
TITLE A collection of 1814 human chromosome 7-specific STS
JOURNAL Genome Res. 7 (1), 59-64 (1997)
MEDLINE 97189344
PUBMED 9037602
REFERENCE 2 (bases 1 to 6128)
AUTHORS Green, E.D.
TITLE Human chromosome 7 STS (1997)
JOURNAL Unpublished (1997)
COMMENT Synonyms: IGFBP1
GDB: GDB:3754042
GDB: DSEG: IGFBP1
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@hgrl.nih.gov
Primer A: TATCAGCAGACAGAGTGG
Primer B: TGAGGACCCAGATCCAG
STS size: 340
PCR Profile:

Presnak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 1.00 minute(s)
Annealing: 60 degrees C for 2.00 minute(s)
Polymerization: 72 degrees C for 2.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer TC

Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 5 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

FEATURES
source
1..6128
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"

This STS was developed from sequence determined by another investigator. See GenBank record M4567 for additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

gene 1..6128
STS 4155..4494
primer_bind /gene="IGFBP1"
primer_bind 4155..4174
primer_bind /gene="IGFBP1"
primer_bind complement(4475..4494)

ORIGIN
Query Match 17.2%; Score 63.8; DB 11; Length 6128;
Best Local Similarity 63.2%; Pred. No. 1.2e-07;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

120 TCCAGGAACACGAGAGTCCCGCCGCGCCATGTCACAGAGAAACAACTATTATTA 179
438 TCTTCCACACGAGGTTGGCTGAGGCGCTTGGGTGACAGCAAACTATTATTGA 497
180 ACACGGGATCTTACGACGCTGCGCTGACATCATTAACCC-----GTGTCGCG 229
498 ACACCTAGCTCTTACGCTGCGCGCTGCGCATCATTAACCTCTGTCAGAGTGGCGG 557
230 AGCCAGCCCTTCAATTAAGCCCTGAGTATGCGCAGCAGATGATGCTGCGCGGAGA 289
558 CCTGTGCCCTTATTAAGTGCGCGCGCTGTGTCCAGCAGATGCGCAGCCATCC---- 613
290 CACAAACCCAGGAGATGTAACACTGC-ACACGCGCATCTGCCCCAGAGCTGTGACA 348
614 ---CATCCAGGAGATGTGCGCGCGCGCGCCGACCTCCAGAGAGACTGGCCA 669
349 CCACCTTC 355
670 CCGCTCC 676

RESULT 15
LOCUS HUMIGFBP1 6480 bp DNA linear PRI 08-NOV-1994
DEFINITION Human insulin-like growth factor binding protein-1 (IGFBP1) gene, complete cds.
ACCESSION M59316 J055683
VERSION M59316.1 GI:184809
KEYWORDS insulin-like growth factor binding protein 1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Cubbage, M.L., Suwanichkul, A. and Powell, D.R.
TITLE Structure of the human chromosome gene for the 25 kilodalton insulin-like growth factor binding protein
JOURNAL Mol. Endocrinol. 3 (5), 846-851 (1989)
MEDLINE 89330502
PUBMED 2474129
REFERENCE 2 (bases 1 to 6480)
AUTHORS Suwanichkul, A., Cubbage, M.L. and Powell, D.R.
TITLE The promoter of the human gene for insulin-like growth factor binding protein-1. Basal promoter activity in HEP 2 cells depends upon liver factor B1
JOURNAL J. Biol. Chem. 265 (34), 21185-21193 (1990)
MEDLINE 91065933
PUBMED 1701175

COMMENT Original source text: Human leukocyte DNA.
FEATURES
source
1..6480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="7p13-p12"
/cell_type="leukocyte"
/gene="IGFBP1"
/join(1206..1719,3258..3427,4643..4771,5662..6362)
/join(1206..1719,3258..3427,4643..4771,5662..6362)
/gene="IGFBP1"

```
exon /note="G00-120-075; putative"
1206. .1719
/gene="IGFBP1"
/note="G00-120-075; putative"
/number=1
join(1371. .1445,1446. .1719,3258. .3427,4643. .4771,
5662. .5793)
/gene="IGFBP1"
/note="putative"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AA52783.1"
/db_xref="GI:184810"
/db_xref="GDB:600-120-075"
/translation="MSRVVARVWLVLTLTVQVGTAGAPWQCAPCSAEKIALCPV
SACSEVTRSGAGCCPMCALPLGAACVATRCARGLSCRALPGEOQLHALTRGQG
ACVQSDASAPHAAGSPESPSTEIEELDFHLMAPSEEDHSILMDAISTYDG
SKALHTNIXKMKERCIETLYRVVSLAKAQETSGEETSKFYLPRCNKNGFYHSRQCE
TSMDEAGLCWCVIPMNGKRIIGSPFIRGDRNCQIYFNVQN"
1371. .1445
/gene="IGFBP1"
/note="G00-120-075"
/evidence=experimental
join(1446. .1719,3258. .3427,4643. .4771,5662. .5790)
/gene="IGFBP1"
/product="insulin-like growth factor binding protein-1"
/note="G00-120-075; putative"
3258. .3427
/gene="IGFBP1"
/note="G00-120-075; putative"
/number=2
4643. .4771
/gene="IGFBP1"
/note="G00-120-075; putative"
/number=3
5662. .6362
/gene="IGFBP1"
/note="G00-120-075; putative"
/number=4

exon
exon
exon
exon

ORIGIN Chromosome 7.
Query Match 17.2%; Score 63.8; DB 9; Length 6480;
Best Local Similarity 63.2%; Pred. No. 1.2e-07;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCAGAGAACACGGAGTGCCTCCGCGCCCATGTCAAGCAAAACACTTATTGGA 179
DB 1047 TCCTCCACACGCGTTTGGTGAAGGCTTGSGTGCACTAGCAAAACACTTATTGGA 1106
QY 180 ACACGGGATCTTACACGCTGCCTGAACATCATTAAACC-----GTGCTGCCG 229
DB 1107 ACATCAGCTCTTACGCTGCCTGCCTGAACATCATTAACTCTGTCGAAGTGCGCGG 1166
QY 230 AGCCAGCCCTTATTAAGGCTCCGGGTATGGCCAGACGATGTCACACTGCCCCGAGAGA 289
DB 1167 CCTGTCCCTTATTAAGGCTCCGGGTATGGCCAGACGATGTCACACTGCCCCGATCC--- 1222
QY 290 CACAAACCCAGGAGCATTTGAACACTGC-ACACGGCATCTGCCCCAGAGAGCTGACCA 348
DB 1223 ----CATCAGAGGACATCTGCGCGCGCGCCGCCCACTCCCAAGAGACACTGSCCA 1278
QY 349 CCACTTC 355
DB 1279 CCGCTCC 1285
```

Search completed: June 7, 2004, 13:58:20
Job time : 2647 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:46:36 ; Search time 339 Seconds
(without alignments)
4661.739 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
1 tacactcggggccagagtc.....ttccgctactagctagccgc 372

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_29Jan04:*

1: genebegn1808:*\n2: genebegn1908:*\n3: genebegn20008:*\n4: genebegn2001as:*\n5: genebegn2001bs:*\n6: genebegn2002s:*\n7: genebegn2003as:*\n8: genebegn2003bs:*\n9: genebegn2003cs:*\n10: genebegn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372	100.0	372	6	ABX15375 Rat insul
2	321	86.3	321	6	ABX15374 Rat insul
3	223.6	60.1	423	6	ABX15376 Rat insul
4	219.4	59.0	270	6	ABX15373 Rat insul
5	219	58.9	219	6	ABX15372 Rat insul
6	154	41.4	423	6	ABX15376 Rat insul
7	76.4	20.5	1500	6	ABK63701 Rat seque
8	76.4	20.5	1500	7	ABT41911 Toxicity
9	76.4	20.5	1500	7	ABT41911 Toxicity
10	76.4	20.5	1500	9	ADBS8201 Toxicity
11	76.4	20.5	5001	4	AAH22429 Rat insul
12	63.8	17.2	6128	6	ABN5886 Gene #239
13	63.8	17.2	6128	7	ABV75371 Human IGF
14	63.8	17.2	6134	4	AAH57489 Human liv
15	55.2	14.8	194	2	AAAT3002 Rat type
16	55.2	14.8	13011	2	AAAT96631 cDNA enco
17	55.2	14.8	13011	7	ABT42448 Toxicity
18	51	13.7	51	6	ABX15380 Rat liver
19	51	13.7	51	6	ABX15371 Rat liver
20	51	13.7	270	6	ABX15373 Rat liver
21	48	12.9	48	6	ABX15379 Rat liver
22	38.6	10.4	44861	6	AA520000 DNA enco
23	37.8	10.2	7061	6	AB132997 Human imm

C 24	37.8	10.2	7061	6	AB170248	Ab170248 Chemical
C 25	37.8	10.2	7061	6	AA661189	AA661189 Human gen
C 26	37.6	10.1	1254	5	AA577024	AA577024 DNA enco
C 27	37.4	10.1	12850	5	ADB36326	ADB36326 Human fac
C 28	36.8	9.9	761	6	AB190266	AB190266 Human pol
C 29	36.6	9.8	1678	6	ADN52441	ADN52441 Human cod
C 30	36.6	9.8	3054	6	AB570481	AB570481 Human bon
C 31	35.8	9.6	11009	6	ABQ72907	ABQ72907 Mouse lam
C 32	35.8	9.6	11009	6	AA170816	AA170816 Mouse lam
C 33	35.6	9.6	490	8	ACH34858	ACH34858 Human end
C 34	35	9.4	3895	6	ABN87725	ABN87725 Human pro
C 35	34	9.1	10537	4	AA535769	AA535769 Human car
C 36	34	9.1	10537	4	AAK69582	AAK69582 Human imm
C 37	34	9.1	10537	9	AD546463	AD546463 Human car
C 38	34	9.1	10543	4	AA535770	AA535770 Human car
C 39	34	9.1	10543	4	AAK69583	AAK69583 Human imm
C 40	34	9.1	10543	9	AD546464	AD546464 Human car
C 41	34	9.1	26277	4	AAK70428	AAK70428 Human imm
C 42	33.4	9.0	24873	8	ADN02594	ADN02594 Human BMI
C 43	33.4	9.0	24873	9	AD572332	AD572332 Human BMI
C 44	33.4	9.0	28000	9	ADD00956	ADD00956 Human Jsg
C 45	33.2	8.9	650	7	AB252705	AB252705 Aspergill

ALIGNMENTS

RESULT 1	ABX15375	ABX15375 standard; DNA; 372 BP.
XX	AC	ABX15375;
XX	DT	17-APR-2003 (first entry)
XX	DE	Rat insulin regulator construct DNA #3.
XX	KW	Glucose response element; GIRE; liver pyruvate kinase; LFK; ds; IGF-1;
XX	KW	insulin-sensitive element; ISE; basal promoter; hyperglycemia; insulin;
XX	KW	insulin-like growth factor binding protein-1; hypoglycemia; glucose;
XX	KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX	KW	glucagon; euglycemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX	KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX	KW	intracellular hormone receptor; insulin regulator construct; anabolic.
XX	OS	Rattus norvegicus.
XX	OS	Synthetic.
XX	PN	US2002107198-A1.
XX	PD	08-AUG-2002.
XX	PF	10-OCT-2001; 2001US-00972916.
XX	PR	11-OCT-2000; 2000US-0239113P.
XX	PI	(THUL/) THUL P M.
XX	PI	Thule PM;
XX	DR	WPI; 2002-674190/72.
XX	PT	New insulin regulator cassette, useful e.g. for treating diabetes,
XX	PT	provides specific, glucose-inducible transgenic expression of insulin in
XX	PT	liver cells.
XX	PS	Claim 9; Page 14; 37pp; English.
XX	CC	The invention relates to an insulin regulator construct comprising a
XX	CC	glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX	CC	promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX	CC	factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX	CC	to treat or prevent diabetic complications, to regulate insulin

CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
CC
SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 372; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 4.9e-103;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACACTGGGGGCGCAGAGTCACAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGG 60
DB 1 TACACTGGGGGCGCAGAGTCACAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGG 60
QY 61 GGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGGCGCCAGAGT 120
DB 61 GGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGGCGCCAGAGT 120
QY 121 CCAGAACCAACGGAGTGGCCCGGCGCCCATGTACAGCAAAACAACTTATTTTAA 180
DB 121 CCAGAACCAACGGAGTGGCCCGGCGCCCATGTACAGCAAAACAACTTATTTTAA 180
QY 181 CACGGGAGTCTTACAGACCGCTGCTGACATCATTAACCCGTCGCGAGCCAGCCCTT 240
DB 181 CACGGGAGTCTTACAGACCGCTGCTGACATCATTAACCCGTCGCGAGCCAGCCCTT 240
QY 241 CATTAAGCCCTTGGTATGAGCCAGACGATGCTCACTGCGCGGAGACCAAAACCCAG 300
DB 241 CATTAAGCCCTTGGTATGAGCCAGACGATGCTCACTGCGCGGAGACCAAAACCCAG 300
QY 301 CGAGCATTTGAACACTGACACAGCGCCATCTGCCAGAGAGCTGTGACCACTTCCTGCTA 360
DB 301 CGAGCATTTGAACACTGACACAGCGCCATCTGCCAGAGAGCTGTGACCACTTCCTGCTA 360
QY 361 CTAGCTAGCCGC 372
DB 361 CTAGCTAGCCGC 372

RESULT 2
ABX15374
ID ABX15374 standard; DNA; 321 BP.

XX ABX15374;

XX 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #2.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
OS Synthetic.

XX US2002107198-A1.

PD 08-AUG-2002.
XX 10-OCT-2001; 2001US-00972916.
XX 11-OCT-2000; 2000US-0239113P.
FR (THUL/) THULE P M.
PA
XX Thule PM;
XX WPI; 2002-674190/72.
DR
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 9; Page 14; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
CC
XX

Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 86.3%; Score 321; DB 6; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.5e-87;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TACACTGGGGGCGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGG 111
DB 1 TACACTGGGGGCGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACACTGGG 60
QY 112 GGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACAGCAAAACAACT 171
DB 61 GGCAGAGTCCAGAACCAACGGAGTGGCCCGGCGCCCATGTACAGCAAAACAACT 171
QY 172 TATTTGAACAGGGAGATCTTACAGACGCTGCTGACATCATTAACCCGTCGCGAG 231
DB 121 TATTTGAACAGGGAGATCTTACAGACGCTGCTGACATCATTAACCCGTCGCGAG 180
QY 232 CGAGCCCTTCAATAAGGCTGTATGAGCCAGACGATGCTCACTGCGCGGAGACA 291
DB 181 CGAGCCCTTCAATAAGGCTGTATGAGCCAGACGATGCTCACTGCGCGGAGACA 240
QY 292 CAACCCAGAGAGATTTGAACACTGACACAGCGCCATCTGCCAGAGAGCTGTACCA 351
DB 241 CAACCCAGAGAGATTTGAACACTGACACAGCGCCATCTGCCAGAGAGCTGTACCA 300
QY 352 CTTCGCTACTAGTACCCGC 372
DB 301 CTTCGCTACTAGTACCCGC 321

RESULT 3
ABX15376
ID ABX15376 standard; DNA; 423 BP.

XX

AC ABX15376;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #4.
XX
KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KM insulin-sensitive element; ISB; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR MPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes.
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis.
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
Query Match 60.1%; Score 223.6; DB 6; Length 423;
Best Local Similarity 76.2%; Pred. No. 6,3e-58;
Matches 227; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

DB 154 CATGGCGACGCGGACCTCCGTGGTCTGTGACTGTGCCCCCAGTATACAGAGA 213
QY 163 AAACAACTTATTTTGAACACGCGGATCTTAGAGCGTCCTGCAATCATTAACCGT 222
DB 214 AAACAACTTATTTTGAACACGCGGATCTTAGAGCGTCCTGCAATCATTAACCGT 273
QY 223 GCTGCGAGCAGCAGCCCTTATTAAGCCCTGGATATGCGCAGCAGCATGTCCACTGCC 282
DB 274 GCTGCGAGCAGCAGCCCTTATTAAGCCCTGGATATGCGCAGCAGCATGTCCACTGCC 333
QY 283 GCGGAGACAAACCCAGGAGATTGAACACTGCACACGCGCATGTGCCAGAGCTG 342
DB 334 GCGGAGACAAACCCAGGAGATTGAACACTGCACACGCGCATGTGCCAGAGAGCTG 393
QY 343 TGACGACCACTTCCGCTACTAGTGGCGG 372
DB 394 TGACGACCACTTCCGCTACTAGTGGCGG 423
RESULT 4
ABX15373
ID ABX15373 standard; DNA; 270 BP.
XX
AC ABX15373;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KM insulin-sensitive element; ISB; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR MPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes.
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 13-14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties

CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 59.0%; Score 219.4; DB 6; Length 270;
Best Local Similarity 99.5%; Pred. No. 1e-56; Indels 0; Gaps 0;
Matches 220; Conservative 0; Mismatches 1;

QY 152 TGTCAAGCAAAACAACTTATTTGAACACGGGGATCTAGACGCTGCGCTGACAT 211
DB 50 TATCAAGCAAAACAACTTATTTGAACACGGGGATCTAGACGCTGCGCTGACAT 109
QY 212 CATTACCCGCTGCTCCGAGACCAAGCCCTTCAATAGGCCCTGGTATGSCCAGCCAGCATG 271
DB 110 CATTACCCGCTGCTCCGAGACCAAGCCCTTCAATAGGCCCTGGTATGSCCAGCCAGCATG 169
QY 272 GTCCACCTGCGCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 331
DB 170 GTCCACCTGCGCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 229
QY 332 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 372
DB 230 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270

RESULT 5
ABX15372
ID ABX15372 standard; DNA; 219 BP.
AC ABX15372;
DT 17-APR-2003 (first entry)
DE Rat insulin-sensitive element (ISE) DNA.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX Rattus norvegicus.
OS
XX US2002107198-A1.
PN
XX 08-AUG-2002.
PD
XX 10-OCT-2001; 2001US-00972916.
PF
XX 11-OCT-2000; 2000US-0239113P.
PR
XX (THUL/) THUL P M.
PA
XX Thule PM;
FI
XX WPI; 2002-674190/72.
DR
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 8; Page 13; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth

CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention
SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 58.9%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e-56; Indels 0; Gaps 0;
Matches 219; Conservative 0; Mismatches 0;

QY 154 TCACAGCAAAACAACTTATTTGAACACGGGGATCTAGACGCTGCGCTGACATCA 213
DB 1 TCACAGCAAAACAACTTATTTGAACACGGGGATCTAGACGCTGCGCTGACATCA 60
QY 214 TTAACCCGCTGCTCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 273
DB 61 TTAACCCGCTGCTCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 120
QY 274 CCACCTGCGCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 333
DB 121 CCACCTGCGCCGAGACCAACCCAGCGAGCATTTGAACCTGACACCGCCATCTGC 180
QY 334 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 372
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219

RESULT 6
ABX15376/C
ID ABX15376 standard; DNA; 423 BP.
AC ABX15376;
DT 17-APR-2003 (first entry)
DE Rat insulin regulator construct DNA #4.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX Rattus norvegicus.
OS
XX Synthetic.
OS
XX US2002107198-A1.
PN
XX 08-AUG-2002.
PD
XX 10-OCT-2001; 2001US-00972916.
PF
XX 11-OCT-2000; 2000US-0239113P.
PR
XX (THUL/) THUL P M.
PA
XX Thule PM;
FI
XX

DR WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPR) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
SQ
Query Match 41.4%; Score 154; DB 6; Length 423;
Best Local Similarity 100.0%; Pred. No. 8, 9e-37;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TACACTGGGGGCCAGAGTCCAGAACCGGAGTGGCCCTGCGCCCATGTACACTGAG 60
DB 204 TACACTGGGGGCCAGAGTCCAGAACCGGAGTGGCCCTGCGCCCATGTACACTGAG 145
QY 61 GGGCAAGTCCAGAACCGGAGTGGCCCTGCGCCCATGTACACTGAGGCGCCAGAT 120
DB 144 GGGCAAGTCCAGAACCGGAGTGGCCCTGCGCCCATGTACACTGAGGCGCCAGAT 85
QY 121 CCAGAACCGGAGTGGCCCTGCGCCCATGT 154
DB 84 CCAGAACCGGAGTGGCCCTGCGCCCATGT 51
RESULT 7
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
XX
XX ABK63701;
XX
XX 18-JUN-2002 (first entry)
XX
XX Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX
XX Rat; 89; hepatotoxin; expressed sequence tag; EST; drug screening;
KW differential expression; centrilobular necrosis; steatosis.
XX
XX Rattus norvegicus.
XX
XX MO200210453-A2.
XX
XX 07-FEB-2002.
XX
XX 30-JUL-2001; 2001WO-US023872.
XX
XX 31-JUL-2000; 2000US-0222040P.
XX
XX 02-NOV-2000; 2000US-0244880P.
XX
XX 11-MAY-2001; 2001US-0290029P.
XX
XX 15-MAY-2001; 2001US-0290645P.
XX
XX 22-MAY-2001; 2001US-0292336P.

PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
PI WPI; 2002-241625/29.
XX
XX WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
PT effects by determining the changes in gene expression in tissues or cells
PT exposed to the toxin and comparing these to gene expression in unexposed
PT tissues or cells.
XX
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
CC compounds or the progression of these toxic effects by determining the
CC global changes in gene expression in tissues or cells exposed to the
CC toxin and comparing these to gene expression in unexposed tissues or
CC cells. Also included are methods of predicting at least one toxic effect
CC of a compound or progression of a toxic effect, preferably the
CC hepatotoxicity of a compound, comprising detecting the level of
CC expression in a tissue or cell sample exposed to the compound of two or
CC more genes listed in the specification, where differential expression of
CC the genes is indicative of at least one toxic effect or progression. The
CC method can also be used to identify an agent which modulates the toxic
CC response and predict cellular pathways that a compound modulates in a
CC cell. The methods utilize a set of at least two probes (on a solid
CC support in kit form), where each of the probes comprises a sequence that
CC specifically hybridises to a gene listed in the specification, a computer
CC system comprising a database containing information identifying the
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC set of genes comprising at least two genes listed in the specification,
CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cells exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterised by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
Query Match 20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 5e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAAACCGAGAGATTTGAACTGCAACGAGGCGCATTTGCCAGAGAGCTGAGCCAC 349
DB 8 CACAAACCGAGAGATTTGAACTGCAACGAGGCGCATTTGCCAGAGAGCTGAGCCAC 67
QY 350 CACTTCGGCTACTAGCTA 367
DB 68 CACTTCGGCTACTATCTA 85
RESULT 8
ABT41911
ID ABT41911 standard; DNA; 1500 BP.
XX
XX ABT41911;
XX
XX 26-JUN-2003 (first entry)
XX
XX Toxicity modelling related rat gene SEQ ID No 1613.
XX
XX

XX Toxic effect: gene expression profile; renal toxicity; toxicity marker;
KW database; drug screening; toxicity assay; rat; ds.
XX
XX Rattus norvegicus.
OS
XX MO200295000-A2.
XX
XX
XX PD 28-NOV-2002.
XX
XX PF 22-MAY-2002; 2002MO-US016173.
XX
XX PR 22-MAY-2001; 2001US-0292335P.
XX PR 13-JUN-2001; 2001US-0297523P.
XX PR 19-JUN-2001; 2001US-0298925P.
XX PR 10-JUL-2001; 2001US-0303807P.
XX PR 10-JUL-2001; 2001US-0303808P.
XX PR 10-JUL-2001; 2001US-0303810P.
XX PR 28-AUG-2001; 2001US-0315047P.
XX PR 27-SEP-2001; 2001US-0324928P.
XX PR 22-OCT-2001; 2001US-0330462P.
XX PR 01-NOV-2001; 2001US-0330867P.
XX PR 21-NOV-2001; 2001US-0331805P.
XX PR 06-DEC-2001; 2001US-0336144P.
XX PR 19-DEC-2001; 2001US-0340873P.
XX PR 21-FEB-2002; 2002US-0357842P.
XX PR 21-FEB-2002; 2002US-0357843P.
XX PR 21-FEB-2002; 2002US-0357844P.
XX PR 15-MAR-2002; 2002US-0364134P.
XX PR 08-APR-2002; 2002US-0370144P.
XX PR 08-APR-2002; 2002US-0370206P.
XX PR 08-APR-2002; 2002US-0370247P.
XX PR 17-APR-2002; 2002US-0372794P.
XX PR 21-APR-2002; 2002US-0371679P.
XX
XX
XX PA (GENE-) GENE LOGIC INC.
XX
XX PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-148464/14.
XX
XX
XX PT Predicting at least one toxic effect of a compound, useful for toxicity
XX modeling, comprises preparing a gene expression profile of a tissue or
XX cell sample exposed to the compound, and comparing the gene expression
XX profile to a database.
XX
XX
XX PS Example 4; Page; 446pp; English.
XX
XX
XX CC The invention relates to a novel method of predicting at least one toxic
XX effect of a compound. The method comprises a gene expression profile of a
XX tissue or cell sample exposed to the compound, and comparing the gene
XX expression profile to a database comprising at least part of the data or
XX information given in the specification. The methods are useful for
XX predicting at least one toxic effect of a compound, predicting the
XX progression of a toxic effect of a compound, predicting the renal
XX toxicity of a compound, or identifying toxicity markers in tissues or
XX cells exposed to known renal toxin. The genes are useful as toxicity
XX markers in drug screening and toxicity assays, in monitoring disease or
XX physiological states, or disease progression. This polynucleotide
XX represents a rat DNA sequence relating to the toxic effect database
XX described in the specification. NOTE: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the World Intellectual Property
XX Organization
XX
XX
XX SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX
XX Query Match 20.5%; Score 76.4; DB 7; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 5e-11; Indels 0; Gaps 0
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Db	8	CACAAACCCGAGCGAGCATTTGAACAACCTGTGACACAGCGGCATCTGTCCGACAGAGCTGTGACAC	67
Qy	350	CACTTCGCGCTACTAGCTA	367
Db	68	CAC TTCGCGCTACTACTA	85
RESULT 9			
ADBS8201			
ID	ADBS8201	standard; DNA; 1500 BP.	
XX			
AC	ADBS8201;		
XX			
DT	04-DEC-2003	(first entry)	
XX			
DE	Toxicity-related gene, SEQ ID 3227.		
XX			
KM	Toxic; toxin; gene expression profile; hepatotoxicity; liver;		
XX	drug screening; toxicity assay; ds.		
OS	Unidentified.		
XX			
PN	WO2003064624-A2.		
XX			
PF	31-JAN-2003; 2003WO-US003194.		
XX			
PR	31-JAN-2002; 2002US-00060087.		
XX			
PR	15-MAR-2002; 2002US-0364045P.		
XX			
PR	15-MAR-2002; 2002US-0364055P.		
XX			
PR	30-DEC-2002; 2002US-0436643P.		
XX			
PA	(GENE-) GENE LOGIC INC.		
XX			
PI	Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;		
XX			
DR	WPI; 2003-689530/65.		
XX			
PT	Predicting a toxic effect of a compound, useful in identifying toxicity		
XX	markers in liver tissues or cells for drug screening and toxicity assays,		
PT	comprises preparing gene expression profile of tissue or cells exposed to		
XX	the compound.		
XX			
PS	Claim 1; SEQ ID NO 3227; 1156bp; English.		
XX			
XX			
CC	The present invention relates to a method for predicting a toxic effect		
CC	of a compound. The method comprises preparing a gene expression profile		
CC	of a tissue or cell sample exposed to the compound, and comparing the		
CC	gene expression profile to a database comprising SEQ ID 1-4925, where		
CC	differential expression of the gene indicates at least one toxic effect.		
CC	The method is useful for predicting at least one toxic effect of a		
CC	compound, predicting hepatotoxicity or the progression of a toxic effect		
CC	of a compound, identifying an agent that modulates the onset or		
CC	progression of a toxic response, predicting the cellular pathways that a		
CC	compound modulates in a cell, and identifying an agent that modulates at		
CC	least one activity of a protein. The method and compositions of the		
CC	present invention using a database of genes having liver toxin-induced		
CC	differential expression, are useful in identifying toxicity markers in		
CC	liver tissues or cells for drug screening and toxicity assays. Note: The		
CC	sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX			
XX			
SQ	Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;		
	Query Match	20.5%; Score 76.4; DB 9; Length 1500;	
	Best Local Similarity	98.7%; Pred. No. 5e-13;	
Db	Matches	77; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
Qy	290	CACAAACCCGAGCGAGCATTTGAACAACCTGTGACACAGCGGCATCTGTCCGACAGAGCTGTGACAC	349
Db	8	CACAAACCCGAGCGAGCATTTGAACAACCTGTGACACAGCGGCATCTGTCCGACAGAGCTGTGACAC	67

OY 350 CACTTCGGCTACTAGCTA 367
DB 68 CACTTCGGCTACTATCTA 85

RESULT 10
ADBS2710
ID ADBS2710 standard; DNA; 1500 BP.

AC ADBS2710;
XX
DT 04-DEC-2003 (first entry)

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Caastle A, Orr M;
PI Elashoff M;
XX
DR WPI; 2003-731472/59.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX
PS Claim 44; SEQ ID NO 3252; 874pp; English.
XX
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC gene listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for

CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
Query Match 20.5%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 5e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 290 CACAAACCGAGCGAGCATTTGAACACTGCAACGGCCATTTGCCACAGAGCTGTGACAC 349
DB 8 CACAAACCGAGCGAGCATTTGAACACTGCAACGGCCATTTGCCACAGAGCTGTGACAC 67
XX
OY 350 CACTTCGGCTACTAGCTA 367
DB 68 CACTTCGGCTACTATCTA 85
XX
RESULT 11
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
XX
AC AAH22429;
XX
DT 22-AUG-2001 (first entry)
XX
DE Rat insulin-like growth factor binding protein nucleotide sequence.
XX
KW Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW non-steroidal antiinflammatory drug; ds.
XX
OS Rattus norvegicus.
XX
PN WO200138579-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032049.
XX
PR 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gould-Rothberg BE, Dipippo VA, Ramseh TM, Gerwein RW;
PI WPI; 2001-355948/37.
XX
DR
XX
PT Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
XX
PS Disclosure; Page 22-24; 76pp; English.
XX
XX
CC The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in a reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX
XX
Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

Query Match 20.5%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 7.3e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 230 CACAAACCCAGGAGCATTTGAACACTGCAACAGGCGCATCTGCCGAGAGCTGTACAC 349
DB 1 CACAAACCCAGGAGCATTTGAACACTGCAACAGGCGCATCTGCCGAGAGCTGTACAC 60
QY 350 CACTTCGCTACTAGCTA 367
DB 61 CACTTCGCTACTAGCTA 78

RESULT 12
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX AC ABN95896;
XX DT 13-AUG-2002 (first entry)
XX DE Gene #2394 used to diagnose liver cancer.
KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumour; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX Homo sapiens.
OS
XX WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US030589.
XX PR 02-OCT-2000; 2000US-0237054P.
XX PA (GENE-) GENE LOGIC INC.
XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX DR
XX PT Diagnosing and detecting the progression of liver cancer, hepatocellular
XX carcinoma or metastatic liver tumor in a patient, involves detecting the
XX level of expression of two or more genes in a liver tissue sample.
XX BS Claim 1; SEQ ID NO 2394; 298bp; English.

CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;

Query Match 17.2%; Score 63.8; DB 6; Length 6128;
Best Local Similarity 63.2%; Pred. No. 5.2e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCACGAAACACGAGGAGTGGCCCGCCCATGTCAACAGCAAAACAACTATTTTGA 179

DB 438 TCTTCCCAACGAGCGTTTGGCTAGAGGCTTGGGTGCTACACAAACAACTATTTTGA 497
QY 180 ACAGGGAGATCTTGAACAGCTGCGCTGACATCAATTAAACC-----GTGCTGCCG 229
DB 498 AACTCACTCTTACCGTGGCGCGCTGCAATCACTTACTCTCTGTGCAAGTGGCGCG 557
QY 230 AGCAGCCCTTCATTAAGGCGCTGGGTATGAGCCAGACAGCATGTGCTACCTGCCCGAGA 289
DB 558 CCGTGGCCCTTATATAGGTGGCGCTGTGTCAAGAGCATCGGCGACCGCCATCC---- 613
QY 290 CACAAACCCAGGAGCATTTGAACACTGC-AACGCGCATCTGCCCGAGAGCTGTACCA 348
DB 614 ----CATCCAGCAGCATCTGCGCGCGCGCCGCCACCTCCAGAGACACTGCGCA 669
QY 349 CCACTTC 355
DB 670 CCGCTCC 676

RESULT 13
ABV75371
ID ABV75371 standard; DNA; 6128 BP.
XX AC ABV75371;
XX DT 07-MAR-2003 (first entry)
XX DE Human IGFBP-1 gene sequence.
XX KM Insulin-like growth factor binding protein; IGFBP; cytosolic; liver;
XX cancer; human; IGFBP-1; gene; ds.
XX OS Homo sapiens.
XX FH Location/Qualifiers
FT CDS
FT 762..5200
FT /tag= a
FT /product= "IGFBP-1"
FT /note= "Insulin-like growth factor binding protein;
FT contains introns"
FT 762..1110
FT /tag= b
FT /number= 1
FT 1111..2656
FT /tag= c
FT /number= 1
FT 2657..2826
FT /tag= d
FT /number= 2
FT 2827..4040
FT /tag= e
FT /number= 2
FT 4041..4169
FT /tag= f
FT /number= 3
FT 4170..5068
FT /tag= g
FT /number= 3
FT 5069..5197
FT /tag= h
FT /number= 4
XX MO200290580-A1.
XX PD 14-NOV-2002.
XX PP 03-MAY-2002; 2002WO-AU000558.
XX PR 03-MAY-2001; 2001US-0288441P.
XX PA (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
PA (SICE-) SINGAPORE GEN HOSPITAL PTE LTD.

PA (ARGA/) ARGAET V P.
XX
PI Huynh TH, Chow PKH, Soo KC;
XX
DR WPI, 2003-103522/09.
DR P-PSDB; ABB82757.
XX
PT Detecting the presence or diagnosing the risk of a liver cancer in a
PT patient comprises detecting aberrant expression of a gene encoding an
PT insulin-like growth factor binding protein.
XX
PS Example; Page 104-108; 142pp; English.
XX
CC The invention relates to detecting the presence or diagnosing the risk of
CC a liver cancer in a patient. The method involves detecting in a
CC biological sample obtained from the patient aberrant expression of a gene
CC encoding an insulin-like growth factor binding protein (IGFBP). The
CC method is useful for detecting the presence or diagnosing the risk of a
CC liver cancer or for screening agents in a patient. The agent is useful
CC for the manufacture of a medicament for treating and/or preventing liver
CC cancer. The present sequence represents a human IGFBP-1 polypeptide
CC encoding genomic DNA (GenBank Accession No. M74587)
XX
SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
XX
Query Match 17.2%; Score 63.8; DB 7; Length 6128;
Best Local Similarity 63.2%; Pred. No. 5.2e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
XX
QY 120 TCCAGAACCCAGGAGTGCCCGCTGCGCCCATGTCAACAAGCAAACTATTATTGA 179
DB 438 TCTTCCACACGAGGTTTGCGTAGGGCTTGCGCATGCAAAACAACTATTATTGA 497
QY 180 ACACGGGAGATCTTACGACGCTGCTGACATATTATACC-----GTGCTGCCG 229
DB 498 ACACCTAGCTCCCTAGGCTGCGGCGCTGCAATTAATCTCTGTGTCAGAGTGGCGCG 557
QY 230 AGCCAGCCCTTCAATTAAGGCTGCTGATGCGCAGCAGCATGTGCACTGCCCGCGAGA 289
DB 558 CCGTGCCCTTTATTAAGTGCGCGCTGTGTCCAGCAGCATCGGCCACCGCATCC--- 613
QY 290 CACAAACCCAGCAGCATTTGAACACTGC-ACACGGCCATCTGCCAGAGAGCTGTGACCA 348
DB 614 ----CATCCAGCAGCATGTGCGCGCGCGCGCGCCGCACTCCAGAGAGCACTGGCCA 669
QY 349 CCACTTC 355
DB 670 CCGCTCC 676
XX
RESULT 14
AAH57489
ID AAH57489 standard; cDNA; 6134 BP.
XX
AC AAH57489;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human liver cell specific cDNA sequence SEQ ID NO:329.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
KW liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN WO200132927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US030396.
XX

PR 04-NOV-1999; 99US-0163508P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Sornasse T, Seilhammer JU, Watson GA;
XX
DR WPI, 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology.
XX
PS Claim 1; Page 246-248; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by them are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC agents. Expression of (I) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
XX
SQ Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;
XX
Query Match 17.2%; Score 63.8; DB 4; Length 6134;
Best Local Similarity 63.2%; Pred. No. 5.3e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
XX
QY 120 TCCAGAACCCAGGAGTGCCCGCTGCGCCCATGTCAACAAGCAAACTATTATTGA 179
DB 438 TCTTCCACACGAGGTTTGCGTAGGGCTTGCGCATGCAAAACAACTATTATTGA 497
QY 180 ACACGGGAGATCTTACGACGCTGCTGACATATTATACC-----GTGCTGCCG 229
DB 498 ACACCTAGCTCCCTAGGCTGCGGCGCTGCAATTAATCTCTGTGTCAGAGTGGCGCG 557
QY 230 AGCCAGCCCTTCAATTAAGGCTGCTGATGCGCAGCAGCATGTGCACTGCCCGCGAGA 289
DB 558 CCGTGCCCTTTATTAAGTGCGCGCTGTGTCCAGCAGCATCGGCCACCGCATCC--- 613
QY 290 CACAAACCCAGCAGCATTTGAACACTGC-ACACGGCCATCTGCCAGAGAGCTGTGACCA 348
DB 614 ----CATCCAGCAGCATGTGCGCGCGCGCGCGCCGCACTCCAGAGAGCACTGGCCA 669
QY 349 CCACTTC 355
DB 670 CCGCTCC 676
XX
RESULT 15
AAT43002/c
ID AAT43002 standard; DNA; 194 BP.
XX
AC AAT43002;
XX
DT 16-JUL-1997 (first entry)
XX
DE Rat type L pyruvate kinase promoter region.
XX
KW Glucose-inducible; rat; pyruvate kinase type L; L-PK; promoter;
KW hyperglycaemia; diabetes; gene therapy; viral vector; defective virus;
KW ss.
XX
OS Rattus sp.
XX

```

FH Key Location/Qualifiers
FT Promoter 16..39
FT /tag= a
FT /label= L4 element
FT /note= "able to interact with MLTF/USF factor"
FT promoter 40..58
FT /tag= b
FT /label= L3 element
FT /note= "able to interact with HNF 4 factor"
XX
XX
XX WO9632489-A1.
XX PD 17-OCT-1996.
XX
XX
XX 12-APR-1996; 96WO-FR000560.
XX PF
XX 14-APR-1995; 95FR-00004558.
XX PR
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
XX PA
XX
XX PI Chen R, Doiron B, Kahn A;
XX
XX WPI, 1996-477139/47.
XX DR
XX
XX PT Recombinant defective virus contg. heterologous gene under control of
XX PT glucose inducible promoter - and infected mammalian cells, esp. for gene
XX PT therapy of diabetes.
XX PS
XX Claim 4; Page 21; 41pp; French.
XX
XX CC A new defective recombinant virus contains at least one heterologous gene
XX CC under control of an expression signal inducible by glucose or its
XX CC analogues. Preferably, the expression signal is derived from the 183 bp
XX CC region located 5' of the type I pyruvate kinase (I-PK) gene, especially
XX CC having the present sequence. Expression of the recombinant virus can be
XX CC controlled by glucose levels, but insulin production can be stopped
XX CC immediately by administering glucagon. This avoids the risk of
XX CC hypoglycaemia caused by excessive insulin production. Mammalian cells
XX CC infected by the virus can be implanted into liver, spleen, pancreas or
XX CC intestine, to provide insulin secretion at predetermined sites. The virus
XX CC is useful in gene therapy to treat and/or prevent diseases associated
XX CC with hyperglycaemia, particularly diabetes, but more generally can be
XX CC used to provide controlled expression of a wide variety of proteins
XX CC
SQ Sequence 194 BP; 50 A; 59 C; 51 G; 34 T; 0 U; 0 Other;

Query Match 14.8%; Score 55.2; DB 2; Length 194;
Best Local Similarity 88.2%; Pred. No. 7.3e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCGGTGCGCCGATGTACACTGGGGGCGAGTCCAGAACCAAGGAGTGCCTCGG 94
DB 78 TGCCACGAGACCTTTGATCACTGGGGGCGAGTCCAGAACCAAGGAGTGCCTCGG 19
QY 95 CGCCCATG 102
DB 18 CGCCCATG 11

```

Search completed: June 7, 2004, 14:04:12
Job time : 341 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: June 7, 2004, 11:17:56 ; Search time 2491 Seconds

(Without alignments)
4459.547 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 tccactggggccagagtc.....tccgctactagtagcgcg 372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estb:*
2: em_estc:*
3: em_estd:*
4: em_este:*
5: em_estf:*
6: em_estg:*
7: em_esth:*
8: em_esti:*
9: em_estj:*
10: em_estk:*
11: em_estl:*
12: em_estm:*
13: em_estn:*
14: em_esto:*
15: em_estp:*
16: em_estq:*
17: em_estr:*
18: em_ests:*
19: em_estt:*
20: em_estu:*
21: em_estv:*
22: em_estw:*
23: em_estx:*
24: em_esty:*
25: em_estz:*
26: em_estaa:*
27: em_estab:*
28: em_estac:*
29: em_estad:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	26.4	1013	10	BF236974 602026605
2	87.8	23.6	331	13	BY783538 BY783538
3	87.8	23.6	331	13	BY794229 BY794229
4	87.8	23.6	354	13	BY771317 BY771317

Result No.	Score	Query Match	Length	DB ID	Description
5	86.8	23.3	401	9	AT785818
6	86.8	23.3	786	9	AT196314
7	86.8	23.3	706	9	AT196314
8	86.8	23.3	765	9	AT196314
9	86.8	23.3	785	9	AT196314
10	86.8	23.3	799	9	AT196314
11	86.8	23.3	848	9	AT196314
12	86.8	23.3	852	9	AT196314
13	86.8	23.3	852	9	AT196314
14	86.8	23.3	852	9	AT196314
15	86.8	23.3	852	9	AT196314
16	86.8	23.3	852	9	AT196314
17	86.8	23.3	852	9	AT196314
18	86.8	23.3	852	9	AT196314
19	86.8	23.3	852	9	AT196314
20	86.8	23.3	852	9	AT196314
21	86.8	23.3	852	9	AT196314
22	86.8	23.3	852	9	AT196314
23	86.8	23.3	852	9	AT196314
24	86.8	23.3	852	9	AT196314
25	86.8	23.3	852	9	AT196314
26	86.8	23.3	852	9	AT196314
27	86.8	23.3	852	9	AT196314
28	86.8	23.3	852	9	AT196314
29	86.8	23.3	852	9	AT196314
30	86.8	23.3	852	9	AT196314
31	86.8	23.3	852	9	AT196314
32	86.8	23.3	852	9	AT196314
33	86.8	23.3	852	9	AT196314
34	86.8	23.3	852	9	AT196314
35	86.8	23.3	852	9	AT196314
36	86.8	23.3	852	9	AT196314
37	86.8	23.3	852	9	AT196314
38	86.8	23.3	852	9	AT196314
39	86.8	23.3	852	9	AT196314
40	86.8	23.3	852	9	AT196314
41	86.8	23.3	852	9	AT196314
42	86.8	23.3	852	9	AT196314
43	86.8	23.3	852	9	AT196314
44	86.8	23.3	852	9	AT196314
45	86.8	23.3	852	9	AT196314

ALIGNMENTS

RESULT 1
LOCUS BF236974
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:4161889 5',
mRNA sequence.
ACCESSION BF236974
VERSION BF236974.1 GI:11150891
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1013)
NIH-MGC <http://mgc.nci.nih.gov/>.
Mammalian Gene Collection (MGC)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
Place: LLM9443 row: 5 column: 02
High quality sequence stop: 581.

FEATURES
Source

Location/Qualifiers
1. .1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161889"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NCI CGAP L19"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 26.4%; Score 98.2; DB 10; Length 1013;
Best Local Similarity 92.8%; Pred. No. 5.8e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 257 TGCCAGCCAGCATGTCCTCCTCCCGCCGAGACACAAACCCAGGAGCATTTGAACACTG 316
Db 9 TGGGAGCCAGCATGTCCTCCTCCCGCCGAGACACACACCCAGGAGCATTTGAACACTG 68

Qy 317 CACAGGCGCATTCGCCAGAGCTGTGACCCACCACTTCCTGCTACTAGCTA 367
Db 69 CACAGGCGCTGTGCCAGAGCTGTGACCCACCACTTCCTGCTACTAGCTA 119

RESULT 2
BY783538 310 bp mRNA linear EST 10-DEC-2003
LOCUS BY783538 RIKEN full-length enriched, 17.5 days embryo whole body
DEFINITION Mus musculus cDNA clone L930176D05 5', mRNA sequence.
ACCESSION BY783538
VERSION BY783538.1 GI:39710177
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 310)
Carinici, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M.,
Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A.,
Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
Pavan, W., Aldinis, V., Nakagawa, A., Heid, W. A., Iwata, H., Kono, T.,
Nakachi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M.,
Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
Targeting a complex transcripome: the construction of the mouse
full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
22703353
12819125

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for
further details.
Location/Qualifiers
1. .331

FEATURES
Source

Location/Qualifiers
1. .1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161889"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NCI CGAP L19"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.9 kb. Constructed by life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 23.6%; Score 87.8; DB 13; Length 310;
Best Local Similarity 92.9%; Pred. No. 3.9e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 269 ATGTCACCTGCGCCGAGACACAAACCCAGGAGCATTTGAACCTGACACGGCCATC 328
Db 2 ATGTCACCTGCGCCGAGACACACACCCAGGAGCATTTGAACCTGACACGGCCATC 61

Qy 329 TGCCAGAGAGCTGTGACCACTTCCTGCTACTAGCTA 367
Db 62 TGCCAGAGAGCTGTGACCACTTCCTGCTACTAGCTA 100

RESULT 3
BY794229 331 bp mRNA linear EST 10-DEC-2003
LOCUS BY794229 RIKEN full-length enriched, 17.5 days embryo whole body
DEFINITION Mus musculus cDNA clone L930292H24 5', mRNA sequence.
ACCESSION BY794229
VERSION BY794229.1 GI:39720868
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 331)
Carinici, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M.,
Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A.,
Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
Pavan, W., Aldinis, V., Nakagawa, A., Heid, W. A., Iwata, H., Kono, T.,
Nakachi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M.,
Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
Targeting a complex transcripome: the construction of the mouse
full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
22703353
12819125

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for
further details.
Location/Qualifiers
1. .331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"

FEATURES
Source

ORIGIN

/clone="L930292H24"
/cissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"

Query Match 23.6%; Score 87.8; DB 13; Length 331;
Best Local Similarity 92.9%; Pred. No. 4e-15; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 269 ATGGTCCACTGCCCCGGGAGACACAAACCCAGGAGACATTGAACATGACACAGGCGCATC 328
DB 2 ATGGTCCACTGCCCCGGGAGACACACACCCAGGAGACATTGAACATGACACAGGCGCATC 61

QY 329 TGCCGAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 62 TGCCGAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 100

RESULT 4 354 bp mRNA linear EST 10-DEC-2003
BY771317 RIKEN full-length enriched, 17.5 days embryo whole body

DEFINITION Mus musculus cDNA clone L930052G15 5', mRNA sequence.
BY771317
KEYWORDS EST.

ACCESSION BY771317.1 GI:39697955
VERSION
KEYWORDS
SOURCE

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 354)
Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M.,
Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
Sugahara, Y., Saito, R., Oosato, N., Fukuda, S., Sato, K., Matshiki, A.,
Higazane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
Pavan, W., Aldinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T.,
Nakunichi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M.,
Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
Miyamatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

TITLES Targeting a complex transcriptome: the construction of the mouse
full-length cDNA encyclopedia

Genome Res. 13 (6B), 1273-1289 (2003)

JOURNAL MEDLINE
PUBMED 22703353
12819125

COMMENT

Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for
further details.

FEATURES

source Location/Qualifiers
1..354
/organism="Mus musculus"
/mol_type="cDNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930052G15"
/cissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo

ORIGIN

whole body"

Query Match 23.6%; Score 87.8; DB 13; Length 354;
Best Local Similarity 92.9%; Pred. No. 4.1e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 269 ATGGTCCACTGCCCCGGGAGACACAAACCCAGGAGACATTGAACATGACACAGGCGCATC 328
DB 2 ATGGTCCACTGCCCCGGGAGACACACACCCAGGAGACATTGAACATGACACAGGCGCATC 61

QY 329 TGCCGAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 62 TGCCGAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 100

RESULT 5 401 bp mRNA linear EST 02-JUL-1999
A1785818
LOCUS U178H05.Y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:188569 5', similar to gb:U81579 M.musculus mRNA for
insulin-like growth factor binding (MUSE);, mRNA sequence.

DEFINITION A1785818
VERSION A1785818.1 GI:5333534
KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 401)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Ponder, B., Swaller, T., Theising, B., Allen, M., Bowers, Y.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.

TITLES The WashU-NCI Mouse EST Project 1999

JOURNAL Unpublished (1999)

Other ESTs: U178H05.X1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972893

Seg primer: custom primer used
High quality sequence step: 126.

FEATURES

source Location/Qualifiers
1..401
/organism="Mus musculus"
/mol_type="cDNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:188569"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Torgano: liver; Vector: pMT18-F13; Site_1: DraIII
(CACTGTG); Site_2: DraIII (CACTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [GTGTGCTTGTG]; digested
and cloned into distinct DraIII sites of the pMT18-F13
vector (5' site CACTGTG, 3' site CACTGTG). xhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTTAAGCTGG and 3' end
primer CGACTGCACTCAAGCA."

ORIGIN primer CGACCTGACGCTCGAGCACA."

Query Match 23.3%; Score 86.8; DB 9; Length 401;
 Best Local Similarity 92.9%; Pred. No. 8.7e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 270 TGGTCACTGCCCCCGAGACACAAACCAGGACATTGAACACTGACACAGGCCACT 329
 1 TGGTCACTGCCCCCGAGACACACACCAGGACATTGAACACTGACACAGGCCACT 60

DB 330 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 367
 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

OY 330 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 367
 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

RESULT 6
 LOCUS A1196314 480 bp mRNA linear EST 14-OCT-1998
 DEFINITION u171a07.y1 Sugano mouse liver m1a Mus musculus cDNA clone IMAGE:1887828 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerothnathi; Muridae; Murinae; Mus. 1 (bases 1 to 480)
 AUTHOR MARXA, M., HILLIER, L., ALLEN, M., BOWLES, M., DIETRICH, N., DUBUQUE, T., GEISEL, S., KUCABA, T., LACY, M., LE, M., MARTIN, J., MORRIS, M., SCHELLENBERG, K., STEPTOE, M., TAN, F., UNDERWOOD, K., MOORE, B., THEISING, B., WYLLIE, T., LEMON, G., SOARES, B., WALSON, R. and WATERSTON, R.
 TITLE The WashU-HMNI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marxa M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MG1:972152
 Seq primer: custom primer used
 High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver m1a"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGATG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGCGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTGTGCTTAAAGCTGG and 3' end

ORIGIN primer CGACCTGACGCTCGAGCACA."

Query Match 23.3%; Score 86.8; DB 9; Length 480;
 Best Local Similarity 92.9%; Pred. No. 9.5e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 270 TGGTCACTGCCCCCGAGACACAAACCAGGACATTGAACACTGACACAGGCCACT 329
 1 TGGTCACTGCCCCCGAGACACACACCAGGACATTGAACACTGACACAGGCCACT 60

DB 330 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 367
 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

OY 330 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 367
 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTACTA 98

RESULT 7
 LOCUS A1530146 706 bp mRNA linear EST 18-MAR-1999
 DEFINITION u189f09.y1 Sugano mouse liver m1a Mus musculus cDNA clone IMAGE:1889609 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530146 GI:4444281
 VERSION A1530146
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerothnathi; Muridae; Murinae; Mus. 1 (bases 1 to 706)
 AUTHOR MARXA, M., HILLIER, L., KUCABA, T., MARTIN, J., BECK, C., WYLLIE, T., PERSON, B., SWALLER, T., GIBBONS, M., PAGE, D., HARVEY, N., SCHUK, R., RITTER, E., KOHN, S., SHIN, T., JACKSON, Y., CARDENAS, M., MCCANN, R., WATERSTON, R. and WALSON, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marxa M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MG1:973933
 Seq primer: custom primer used
 High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver m1a"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGATG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGCGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTGTGCTTAAAGCTGG and 3' end

ORIGIN

primer CGACCTGCAGCTCGAGACA."

Query Match 23.3%; Score 86.8; DB 9; Length 706;
Best Local Similarity 92.9%; Pred. No. 1.2e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCACTGCCCGCGGAGACACAAACCCAGAGAGATTGAACATGACACAGGCCATCT 329
DB 1 TGGTCACTGCCCGCGGAGACACACACCCAGAGAGATTGAACATGACACAGGCCATCT 60

QY 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98

RESULT 8
A1098594 765 bp mRNA linear EST 20-AUG-1998
LOCUS u911f01.Y1 Sugano mouse liver mla Mus musculus cDNA clone
DEFINITION IMAGE:1881988 5' similar to gb:U81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1098594
VERSION A1098594.1 GI:3448119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 765)

REFERENCE
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacey, M., Le M., Martin, J., Morris, M.,
Schelienberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 395.

FEATURES

source

1..765
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1881988"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMR188-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGTGCTTACTGTG), digested
and cloned into distinct DraIII sites of the pMR188-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGACA."

ORIGIN

Query Match 23.3%; Score 86.8; DB 9; Length 765;
Best Local Similarity 92.9%; Pred. No. 1.2e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCACTGCCCGCGGAGACACAAACCCAGAGAGATTGAACATGACACAGGCCATCT 329
DB 1 TGGTCACTGCCCGCGGAGACACACACCCAGAGAGATTGAACATGACACAGGCCATCT 60

QY 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98

RESULT 9
A1530313 785 bp mRNA linear EST 18-MAR-1999
LOCUS u191f01.Y1 Sugano mouse liver mla Mus musculus cDNA clone
DEFINITION IMAGE:1889785 5' similar to gb:M59316.mal INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:U81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1530313
VERSION A1530313.1 GI:4444448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 785)

REFERENCE
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
MGI:974109
Seq primer: custom primer used
High quality sequence stop: 459.

FEATURES

source

1..785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMR188-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TGTGTGCTTACTGTG), digested
and cloned into distinct DraIII sites of the pMR188-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGCG and 3' end
primer CGACCTGCAGCTCGAGACA."

ORIGIN

Query Match 23.3%; Score 86.8; DB 9; Length 785;
Best Local Similarity 92.9%; Pred. No. 1.2e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGTTCACCTGCCCGCCGAGACACAAACCCAGGACATTGAACACTGACACAGGCGCATCT 329
DB 1 TGTTCACCTGCCCGCCGAGACACACACCCAGGACATTGAACACTGACACAGGCGCATCT 60
QY 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTCTA 98

RESULT 10

LOCUS A1529393 799 bp mRNA linear EST 18-MAR-1999
DEFINITION u187c03.y1 Sugano mouse liver mla Mus musculus cDNA clone

IMAGE:1889392.5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1529393
VERSION A1529393.1 GI:4444074
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 799)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Ponderwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:973716

Seq primer: custom primer used
High quality sequence stop: 506.

FEATURES

source

1..799
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889392"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site: 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TCTTGAGCTCTGCG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end
primer CGACCTGACGCTGAGACCA."

ORIGIN

Query Match 23.3%; Score 86.8; DB 9; Length 799;
Best Local Similarity 92.9%; Pred. No. 1.2e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGTTCACCTGCCCGCCGAGACACAAACCCAGGACATTGAACACTGACACAGGCGCATCT 329
DB 1 TGTTCACCTGCCCGCCGAGACACACACCCAGGACATTGAACACTGACACAGGCGCATCT 60
QY 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTCTA 98

RESULT 11

LOCUS A1790802 848 bp mRNA linear EST 02-JUL-1999
DEFINITION uk28b10.y1 Sugano mouse kidney mla Mus musculus cDNA clone

IMAGE:1970299.5' similar to gb:X81579 M. musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1790802
VERSION A1790802.1 GI:5338518
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 848)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Ponderwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other ESTs: uk28b10.x1

Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:987039

Seq primer: custom primer used
High quality sequence stop: 514.

FEATURES

source

1..848
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse kidney mla"
/note="Organ: kidney; Vector: pME18S-FL3; Site: 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TCTTGAGCTCTGCG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGCG and 3' end
primer CGACCTGACGCTGAGACCA."

ORIGIN

Query Match 23.3%; Score 86.8; DB 9; Length 848;
Best Local Similarity 92.9%; Pred. No. 1.3e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCCACTGCCCCCGAGACACAAACCCAGGAGATTGAACACTGCACACGGCCATCT 329
DB 1 TGGTCCACTGCCCCCGAGAGACACACCCAGGAGATTGAACACTGCACACGGCCGCT 60

QY 330 GCCCAGAGAGCTGTGACCACTTCCGCTACTACTA 367
DB 61 GCCCAGAGAGCTGTGACCACTTCCGCTACTACTA 98

RESULT 12
LOCUS A1528304 852 bp mRNA linear EST 18-MAR-1999
DEFINITION u19510.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1890210 5' similar to gb:M59316.mal INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.
ACCESSION A1528304 GI:4442439
VERSION A1528304
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 852)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,K.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534
Seg primer: custom primer used
High quality sequence stop: 478.
Location/Qualifiers
1..852
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACCATG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG]; digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCATG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGG and 3' end
primer CGACCTGCAGCTGAGACACA."

ORIGIN

Query Match 23.3%; Score 86.8; DB 9; Length 852;
Best Local Similarity 92.9%; Pred. No. 1.3e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCCACTGCCCCCGAGACACAAACCCAGGAGATTGAACACTGCACACGGCCATCT 329
DB 1 TGGTCCACTGCCCCCGAGAGACACACCCAGGAGATTGAACACTGCACACGGCCGCT 60

QY 330 GCCCAGAGAGCTGTGACCACTTCCGCTACTACTA 367
DB 61 GCCCAGAGAGCTGTGACCACTTCCGCTACTACTA 98

RESULT 13
LOCUS A1196154 605 bp mRNA linear EST 14-OCT-1998
DEFINITION u169408.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1887663 5' similar to gb:M59316.mal INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.
ACCESSION A1196154 GI:3748760
VERSION A1196154
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 605)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:971987
Seg primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers
1..605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACCATG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG]; digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CACCATG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTAAAGCTGG and 3' end
primer CGACCTGCAGCTGAGACACA."

Oy 282 CGCCGAGACACAACCAGGAGCATTTGAACACTGACACAGGCGCATCTGCGCCAGAGGCT 341
Db 17 CGAGGAGACACACACCCAGGAGCATTTGAACACTGACACAGGCGCATCTGCGCCAGAGGCT 76
Oy 342 GTGACCCACCACTTCGGCTACTAGCTA 367
Db 77 GTGACCCACCACTTCGGCTACTAGCTA 102

Search completed: June 7, 2004, 14:45:58
Job time : 2495 secs

This Page Blank (uspto)